

# Rotated Eigenstructure Analysis for Source Localization without Energy-decay Models

Junting Chen and Urbashi Mitra

Ming Hsieh Department of Electrical Engineering, University of Southern California  
Los Angeles, CA 90089 USA, email:{juntingc, ubli}@usc.edu

**Abstract**—Herein, the problem of simultaneous localization of two sources given a modest number of samples is examined. In particular, the strategy does not require knowledge of the target signatures of the sources *a priori*, nor does it exploit classical methods based on a particular decay rate of the energy emitted from the sources as a function of range. General structural properties of the signatures such as unimodality are exploited. The algorithm localizes targets based on the rotated eigenstructure of a reconstructed observation matrix. In particular, the optimal rotation can be found by maximizing the ratio of the dominant singular value of the observation matrix over the nuclear norm of the optimally rotated observation matrix. It is shown that this ratio has a unique local maximum leading to computationally efficient search algorithms. Moreover, analytical results are developed to show that the squared localization error decreases at a rate  $n^{-3}$  for a Gaussian field with a single source, where  $n(\log n)^2$  scales proportionally to the number of samples  $M$ .

## I. INTRODUCTION

Underwater source detection and localization is an important but challenging problem. Classical range-based or energy-based source localization algorithms usually require energy-decay models and the knowledge of the environment [1]–[6]. However, critical environment parameters may not be available in many underwater applications, in which case, classical model-dependent methods may break down, even when the measurement signal-to-noise ratio (SNR) is high.

There have been some studies on source localization using nonparametric machine learning techniques, such as kernel regressions and support vector machines [7]–[10]. However, these methods either require a large amount of sensor data, or some implicit information of the environment, such as the choice of kernel functions. For example, determining the best kernel parameters (such as bandwidth) is very difficult given a small amount of data.

This paper focuses on source detection and localization problems when only some structural properties of the energy field generated by the sources are available. Specifically, instead of requiring the knowledge of how energy decays with distance to the source, the paper aims at exploiting only the assumption that the closer to the source the higher energy received, and moreover, the energy field of the source is spatially invariant and decomposable. In fact, such a structural property is generic in many underwater applications. The prior work [11], [12] studied the single source case, where an observation matrix is formed from a few energy measurements of the field in the target area, and the missing entries of the observation

matrix are filled using matrix completion methods. Knowing that the matrix would be rank-1 under full and noise-free sampling of the whole area, singular value decomposition (SVD) is applied to extract the dominant singular vectors, and the source location is inferred from analyzing the peaks of the singular vectors.

Herein, we propose to improve upon two shortcomings in [11], [12]: we make rigorous an estimation/localization bound (versus focusing on the reduction of the search region) and we provide a method for localizing two sources. In the two source case, we need to tackle an additional difficulty that the SVD of the observation matrix does not correspond to the signature vectors of the sources. To resolve this issue, a method of rotated eigenstructure analysis is proposed, where the observation matrix is formed by rotating the coordinate system such that the sources are aligned in a row or in a column of the matrix. We develop algorithms to first localize the central axis of the two sources, and then separate the sources on the central axis.

To summarize, we derive algorithms to simultaneously localize up to two sources based on only a few power measurements in the target area without knowing any specific energy-decay model. The contributions of this paper are as follows:

- We derive the location estimators with analytical results to show that the squared error decreases at a rate  $n^{-3}$  for a Gaussian field with a single source, where  $n(\log n)^2$  scales proportionally to the number of samples  $M$ .
- We develop a localization algorithm for the double source case based on a novel rotated eigenstructure analysis. We show that the two sources can be separated even when their aggregate power field has a single peak.

The rest of the paper is organized as follows. Section II gives the system model and assumptions. Section III develops location estimator with performance analysis for single source case. Section IV proposes rotated eigenstructure analysis for double source case. Numerical results are given in Section V and Section VI concludes this work.

## II. SYSTEM MODEL

Consider that there are  $K$  ( $K = 1, 2$ ) sources with unknown locations  $\mathbf{s}_k = (x_k^S, y_k^S) \in \mathbb{R}^2$  located in a bounded area  $\mathcal{A}$ . Suppose that the sensors can only measure the aggregate power transmitted by the sources, and is given by

$$h(x, y) = \sum_k h_k(x, y)$$

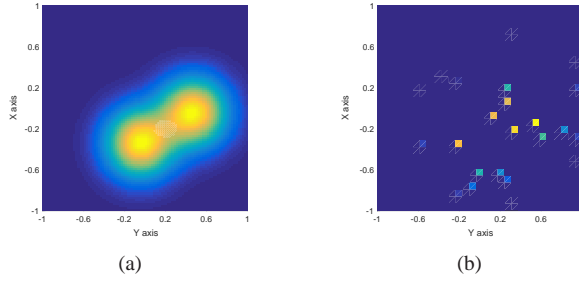


Figure 1. An application example: To localize the two sources in (a) based on the a small number of measurement samples in (b), without knowing the energy-decay models.

for measurement location  $(x, y)$ , where

$$h_k(x, y) = \alpha u(x - x_k^S)u(y - y_k^S) \quad (1)$$

is the power density from source  $k$ , where  $\alpha > 0$ . The explicit form of the density function  $h_k(x, y)$  is unknown to the system, except that the *characteristic* function  $u(x)$  is known to have the following properties

- a) positive semi-definite, i.e.,  $u(x) \geq 0$  for all  $x \in \mathbb{R}$
- b) symmetric, i.e.,  $u(x) = u(-x)$
- c) unimodal, i.e.,  $u'(x) < 0$  for  $x > 0$ ,
- d) smooth, i.e.,  $|u'(x)| < K_u$  for some  $K_u > 0$ , and
- e) normalized, i.e.,  $\int_{-\infty}^{\infty} u(x)^2 dx = 1$ .

Note that  $u(x)$  can be considered as the marginal power density function.

Consider that  $M$  power measurements  $\{h^{(l)}\}$  are taken over distinct locations  $\mathbf{z}^{(l)} = (x^{(l)}, y^{(l)})$ ,  $l = 1, 2, \dots, M$ , uniformly at random in the target area  $\mathcal{A}$ . The measurements are assigned to a  $n_1 \times n_2$  observation matrix  $\hat{\mathbf{H}}$  as follows. First, partition the target area  $\mathcal{A}$  into  $n_1 \times n_2$  disjoint cells  $\mathcal{G}_{ij}$ ,  $i = 1, 2, \dots, n_1$  and  $j = 1, 2, \dots, n_2$ , where  $n_1$  and  $n_2$  are to be determined. Second, assign the power measurements  $h^{(l)}$  to the corresponding  $(i, j)$ th entry of  $\hat{\mathbf{H}}$  as

$$\hat{H}_{ij} = s(\mathcal{G}_{ij})h^{(l)} \quad (2)$$

if  $\mathbf{z}^{(l)} \in \mathcal{G}_{ij}$ , where  $s(\mathcal{G}_{ij})$  measures the area of  $\mathcal{G}_{ij}$ .<sup>1</sup> Denote  $\Omega$  as the set of observed entries of  $\hat{\mathbf{H}}$ , i.e.,  $(i, j) \in \Omega$  if there exists  $\mathbf{z}^{(l)} \in \mathcal{G}_{ij}$  such that  $h^{(l)}$  is assigned to  $\hat{H}_{ij}$ .

For easy discussion, assume that  $\mathcal{A} = [-\frac{L}{2}, \frac{L}{2}] \times [-\frac{L}{2}, \frac{L}{2}]$ ,  $n_1 = n_2 = n$ , and  $\mathcal{G}_{ij}$  are rectangles centered at  $(x_i, y_j)$ ,  $x_i = -\frac{L}{2} + \frac{L}{2n} + \frac{L}{n}(i-1)$ ,  $y_j = -\frac{L}{2} + \frac{L}{2n} + \frac{L}{n}(j-1)$ , and have identical size with each other. Let  $\mathbf{H} = \alpha \sum_{k=1}^K \mathbf{u}_k \mathbf{v}_k^T$  be the matrix of ideal observation, where

$$\mathbf{u}_k = \frac{L}{N} [u(x_1 - x_k^S), u(x_2 - x_k^S), \dots, u(x_n - x_k^S)]^T \quad (3)$$

$$\mathbf{v}_k = \frac{L}{N} [u(y_1 - y_k^S), u(y_2 - y_k^S), \dots, u(y_n - y_k^S)]^T \quad (4)$$

for  $k = 1, 2$ . Thus  $\mathbf{H}$  has rank at most  $K$ . For  $(i, j) \in \Omega$ , we have  $\hat{H}_{ij} \approx H_{ij}$ , where the slight difference is due to sampling away from the centers of the cells  $\mathcal{G}_{ij}$ . As a result,

<sup>1</sup>If multiple samples are close to each other and assigned to the same entry of  $\hat{\mathbf{H}}$ , the value of that entry is the average of the sample values.

$\hat{\mathbf{H}}$  is a sparse and noisy observation of the low rank matrix  $\mathbf{H}$ . An application example is illustrated in 1.

The goal of this paper is to find the approximate locations of the sources using only the spatial invariant property (1) and the four generic properties of the characteristic function  $u(x)$ . Note that this problem is non-trivial. We insist on several features of the algorithm to be developed: it should be robust to structural knowledge of the signatures of the sources (as captured by  $g(x, y)$  in (1)). This disallows the use of parametric regression or parameter estimation for source localization. In addition, we wish to under-sample the target area using small  $M$ . As such, maximum value entries may not represent the true locations of the sources. While not a focus of the current work, we will use matrix completion methods and the low rank property of  $\mathbf{H}$  as in [11], [12] to cope with the under-sampled observations.

### III. EIGENSTRUCTURE ANALYSIS FOR SINGLE SOURCE LOCALIZATION

To simplify the discussion, the following mild assumptions are made.<sup>2</sup>

- A1) The observation area  $\mathcal{A}$  is large enough, such that there is only negligible energy spreading outside the area  $\mathcal{A}$ .
- A2) The parameter  $n$  is not too small, such that  $u(x_i - x_k^S)^2 \delta^2 \approx \int_{x_i}^{x_{i+1}} u(x - x_k^S)^2 dx$  and  $u(y_i - y_k^S)^2 \delta^2 \approx \int_{y_i}^{y_{i+1}} u(y - y_k^S)^2 dy$  for all  $i = 1, 2, \dots, n$ .

Mathematically, the above assumptions imply that the vectors  $\mathbf{u}_k$  and  $\mathbf{v}_k$  have unit norm.

#### A. Observation Matrix Construction

We first exploit the low rank property of  $\mathbf{H}$  to obtain the full matrix  $\hat{\mathbf{H}}_c$  from the partially observed matrix  $\hat{\mathbf{H}}$ . Let  $\mathcal{P}_\Omega(\mathbf{X})$  be a projection, such that the  $(i, j)$ th element of matrix  $\mathcal{P}_\Omega(\mathbf{X})$  is  $[\mathcal{P}_\Omega(\mathbf{X})]_{ij} = X_{ij}$  if  $(i, j) \in \Omega$ , and  $[\mathcal{P}_\Omega(\mathbf{X})]_{ij} = 0$  otherwise. The completed matrix  $\hat{\mathbf{H}}_c$  can be found as the unique solution to the following problem

$$\begin{aligned} & \underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{X}\|_* \\ & \text{subject to} \quad \|\mathcal{P}_\Omega(\mathbf{X} - \hat{\mathbf{H}})\|_F \leq \epsilon \end{aligned} \quad (5)$$

where  $\|\mathbf{X}\|_*$  denotes the nuclear norm of  $\mathbf{X}$  and  $\epsilon$  is a small parameter to tolerate the discrepancy between the two matrices.

To choose a proper dimension  $n$  for the observation matrix  $\hat{\mathbf{H}}_c \in \mathbb{R}^{n \times n}$ , we consider the results in [13]. It has been shown that under some mild conditions of  $\mathbf{H}$  (such as the strong incoherence property and small rank property), the matrix  $\mathbf{H} \in \mathbb{R}^{n \times n}$  can be exactly recovered with a high probability, if the dimension  $n$  satisfies  $Cn(\log n)^2 \leq M$  and noise-free sampling,  $\hat{H}_{ij} = H_{ij}$  for  $(i, j) \in \Omega$ , is performed. Here,  $C$  is a positive constant. Given this, we propose to choose  $n = n_c$  as the largest integer to satisfy  $n_c(\log n_c)^2 \leq M/C$ .

<sup>2</sup>The two assumptions are mainly to avoid discussing the effects on the boundary of  $\mathcal{A}$  and the high order noise term in the sampling noise model (20). Straight-forward modifications can be made to handle the boundary effect in practical algorithms.

### B. Location Estimator Exploiting Property of Symmetry

Consider the SVD of the completed matrix  $\hat{\mathbf{H}}_c$  as  $\hat{\mathbf{H}}_c = \alpha_1 \hat{\mathbf{u}}_1 \hat{\mathbf{v}}_1^T + \sum_{i=2}^{n_c} \alpha_i \hat{\mathbf{u}}_i \hat{\mathbf{v}}_i^T$ . We thus model the singular vectors of  $\hat{\mathbf{H}}_c$  as  $\hat{\mathbf{u}}_1 = \mathbf{u}_1 + \mathbf{e}_u$  and  $\hat{\mathbf{v}}_1 = \mathbf{v}_1 + \mathbf{e}_v$ .

Note that the vectors  $\mathbf{u}_1$  and  $\mathbf{v}_1$  defined in (3) and (4), respectively, contain the source location information due to the unimodal property of  $u(x)$ . However, due to the noise vectors  $\mathbf{e}_u$  and  $\mathbf{e}_v$ , the source location cannot be found by simply locating the peaks of  $\hat{\mathbf{u}}_1$  and  $\hat{\mathbf{v}}_1$ .

To resolve this difficulty, we exploit the symmetric property of  $u(x)$  and develop a location estimator as follows.

Define a reflected correlation function as

$$\hat{R}(t; \hat{\mathbf{u}}_1) = \int_{-\infty}^{\infty} \hat{u}(x) \hat{u}(-x + t) dx \quad (6)$$

where  $\hat{u}(x)$  is a (nonparametric) regression function from vector  $\hat{\mathbf{u}}_1$ . For example,  $\hat{u}(x)$  can be obtained by  $\hat{u}(x) = \hat{\mathbf{u}}_1(i)$  if  $x = x_i$ , and by linear interpolation between  $\hat{\mathbf{u}}_1(i)$  and  $\hat{\mathbf{u}}_1(i+1)$  if  $x_i < x < x_{i+1}$ . Then the location estimator for  $x_1^S$  is given by

$$\hat{x}_1^S(\hat{\mathbf{u}}_1) = \frac{1}{2} \operatorname{argmax}_{t \in \mathbb{R}} \hat{R}(t; \hat{\mathbf{u}}_1). \quad (7)$$

The location estimator for  $y_1^S$  can be obtained in a similar way.

The location estimator (7) exploits the fact that as  $\hat{\mathbf{u}}_1$  is symmetric, the reflected correlation (6), which is the correlation between  $\hat{\mathbf{u}}_1$  and a reflected and shifted version of  $\hat{\mathbf{u}}_1$ , is maximized at the source location. Therefore, the estimator  $\hat{x}_1^S(\hat{\mathbf{u}}_1)$  tries to suppress the perturbation from the noise by correlating over all the entries of  $\hat{\mathbf{u}}_1$ .

We establish several properties for the estimator  $\hat{x}_1^S(\hat{\mathbf{u}}_1)$ .

Consider the autocorrelation for the characteristic function  $u(x)$  as

$$\tau(t) = \int_{-\infty}^{\infty} u(x) u(x - t) dx. \quad (8)$$

Then, the following property can be derived.

**Lemma 1** (Monotonicity). *The autocorrelation function  $\tau(t)$  is non-negative and symmetric. In addition,  $\tau(t)$  is strictly decreasing in  $t > 0$ .*

In addition, the estimator  $\hat{x}_1^S$  is unbiased as summarized in the following proposition.

**Proposition 1** (Unbiased estimator). *Suppose that the elements of the noise vector  $\mathbf{e}_u$  are zero mean and independent and identically distributed (i.i.d.). Consider that linear interpolation is used to construct the regression function  $\hat{u}(x)$  in (6). Under assumptions A1 and A2, the estimator  $\hat{x}_1^S(\hat{\mathbf{u}}_1)$  in (7) is unbiased, i.e.,  $\mathbb{E}\{\hat{x}_1^S(\hat{\mathbf{u}}_1)\} = x_1^S$ .*

Furthermore, the estimation error of  $\hat{x}_1^S$  can be bounded as follows.

**Theorem 1** (Localization error bound). *Consider that  $\epsilon$  in (5) is chosen such that  $\|\mathcal{P}_\Omega(\hat{\mathbf{H}} - \mathbf{H})\|_F \leq \epsilon$ . In addition, suppose that the elements of the noise vector  $\mathbf{e}_u$  are zero mean and i.i.d.. Then, with high probability,*

$$|\hat{x}_1^S - x_1^S| \leq \frac{1}{2} \tau^{-1}(1 - \mu_u L^6 n_c(M)^{-3} + o(n_c(M)^{-3})) \quad (9)$$

where  $\tau^{-1}(r)$  is the inverse function of  $r = \tau(t)$ ,  $\mu_u = 128u(0)^2 K_u^2$ , and  $n_c(M)$  is the largest integer chosen such that  $M \geq C n_c(\log n_c)^2$ .

The specific performance from (9) depends on the characteristics of the energy field. Intuitively, if  $u(x)$  has a sharp peak (large slope of the autocorrelation function  $\tau(t)$ ), the localization error should be smaller. Consider a numerical example where the energy field has a Gaussian characteristic function.

**Corollary 1** (Squared error bound in Gaussian field). *For a Gaussian characteristic function  $u(x) = (\frac{2\alpha}{\pi})^{\frac{1}{4}} e^{-\gamma x^2}$ , there exists a numerical constant  $C_\mu$  such that with high probability, the squared estimation error is upper bounded by*

$$|\hat{x}_1^S - x_1^S|^2 + |\hat{y}_1^S - y_1^S|^2 \leq C_\mu L^6 n_c(M)^{-3} + o(n_c(M)^{-3}). \quad (10)$$

Theorem 1 and Corollary 1 gives the asymptotic performance of the proposed localization algorithm without knowing the energy-decay model. For large  $M$ , the worst case squared error decays at a rate  $n_c(M)^{-3}$ . As a benchmark, the squared error of a naive scheme, which estimates the source location directly from the position of the measurement sample that observes the highest power, decreases as  $M^{-1}$ , which is equivalent to  $n_c(M)^{-1}(\log n_c(M))^{-2}$ , much slower than that of the proposed algorithm. This is because, the granularity of the original observations is  $L/\sqrt{M}$ . The results then confirm that by exploiting the low rank property using matrix completion and the reflected correlation technique, the proposed algorithm significantly improves the localization resolution.

## IV. ROTATED EIGENSTRUCTURE ANALYSIS FOR DOUBLE SOURCE LOCALIZATION

The location estimator  $\hat{x}_1^S$  in (7) is based on the intuition that the singular vectors of  $\mathbf{H}$  are just the vectors  $\mathbf{u}_1$  and  $\mathbf{v}_1$ , which contains the source location in their peaks. However, a similar technique cannot be applied to the two source case, because  $\mathbf{u}_k$  and  $\mathbf{v}_k$  may not be the singular vectors of  $\mathbf{H}$ , as the vectors  $\{\mathbf{u}_k\}$  may not be orthogonal.

### A. Optimal Rotation of the Observation Matrix

When there are two sources, the (ideal) observation matrix  $\mathbf{H}$  is not rank-1, expect for the special case where the two sources are aligned on one of the axes of the coordinate system.

Without loss of generality (w.l.o.g.), assume that the sources are aligned with the  $x$ -axis, where  $y_k^S = C$  for  $k = 1, 2$ . Consequently, we have  $\mathbf{v}_1 = \mathbf{v}_2$ , and  $\mathbf{H} = \alpha(\sum_k \mathbf{u}_k) \mathbf{v}_1^T$ , which is rank-1. Hence, the right singular vector of  $\mathbf{H}$  is  $\mathbf{v}_1$  and, by analyzing the peak of  $\mathbf{v}_1$ , the central axis  $\hat{y}_k^S = C$  can be estimated.

The above observations suggest that we rotate the coordinate system such that the sources are aligned with one of the axes. Consider rotating the coordinate system by  $\theta$ . The entries of  $\hat{\mathbf{H}}_c$  are rearranged into a new observation matrix  $\hat{\mathbf{H}}_\theta$ , where

$$[\hat{\mathbf{H}}_\theta]_{(i,j)} = [\hat{\mathbf{H}}_c]_{(p,q)} \quad (11)$$

in which  $(p, q)$  is the index such that  $(x'_p, y'_q)$  is the closest point in Euclidean distance to  $(\bar{x}, \bar{y})$  in the original coordinate system  $\mathcal{C}_0$ , with  $\bar{x} = d \cos(\beta + \theta)$  and  $\bar{y} = d \sin(\beta + \theta)$ . Here  $\beta = \angle(x_i, y_j)$  is the angle of  $(x_i, y_j)$  to the  $x$ -axis of the rotated coordinate system  $\mathcal{C}_\theta$ , and  $d = \|(x_i, y_j)\|_2$ . Note that  $1 \leq i, j \leq n'$ , where  $n' \leq n_c$ , since the rotation of the axes induce truncation of some data samples.

Let the orientation angle of the central axis of the sources with respect to (w.r.t.) the  $x$ -axis in the original coordinate system  $\mathcal{C}_0$  be  $\theta_0$ ,  $\theta_0 \in [0, \pi)$ . Then the desired rotation for coordinate system  $\mathcal{C}_\theta$  would be  $\theta^* = \theta_0$  for  $\theta_0 < \frac{\pi}{2}$ , or  $\theta^* = \theta_0 - \frac{\pi}{2}$  for  $\theta_0 \geq \frac{\pi}{2}$ . The desired rotation  $\theta$  can be obtained as

$$\underset{\theta \in [0, \frac{\pi}{2}]}{\text{maximize}} \quad \rho(\theta) \triangleq \frac{\lambda_1(\hat{\mathbf{H}}_\theta)}{\sum_k \lambda_k(\hat{\mathbf{H}}_\theta)} \quad (12)$$

where  $\lambda_k(\mathbf{A})$  is the  $k$ th largest singular value of  $\mathbf{A}$ . Note that  $\rho(\theta) \leq 1$  for all  $\theta \in [0, \frac{\pi}{2}]$  and  $\rho(\theta^*) = 1$ , where  $\hat{\mathbf{H}}_\theta$  becomes a rank-1 matrix when the sources are aligned with one of the axes.

The maximization problem (12) is in general non-convex. An exhaustive search for the solution  $\theta^*$  is computationally expensive, since for each  $\theta$ , SVD should be performed to obtain the singular value profile of  $\hat{\mathbf{H}}_\theta$ . Therefore, we need to study the properties of the alignment metric  $\rho(\theta)$  in order to develop efficient algorithms for the source detection.

### B. The Unimodal Property

We also show that the function  $\rho(\theta)$  also has the unimodal property defined as follows.

**Definition 1** (Unimodality). A function  $f(x)$  is called unimodal in a bounded region  $(a, b)$ , if there exists  $x_0 \in [a, b]$ , such that  $f'(x)f'(y) < 0$  for any  $a < x < x_0 < y < b$ .

The unimodality suggests that  $f(x)$  has a single peak in  $(a, b)$ , and hence  $f(x)$  has a unique local maximum (or minimum).

**Theorem 2** (Unimodality in the two source case). *The function  $\rho(\theta)$  in (12) is unimodal in  $\theta \in (\theta^* - \frac{\pi}{4}, \theta^* + \frac{\pi}{4})$ , if*

$$s \cdot \tau'(t) > t \cdot \tau'(s) \quad (13)$$

for all  $0 < s < t$ , where  $\tau'(t) \triangleq \frac{d}{dt}\tau(t)$ . In addition,  $\rho(\theta)$  is strictly increasing over  $(\theta^* - \frac{\pi}{4}, \theta^*)$  and strictly decreasing over  $(\theta^*, \theta^* + \frac{\pi}{4})$ .

The result in Theorem 2 is powerful, since it confirms that the function  $\rho(\theta)$  is unimodal within a  $\frac{\pi}{2}$ -window, and there is a unique local maximum, when the autocorrelation of the energy field characteristic function  $u(x)$  agrees with the condition (13). Note that  $\rho(\theta)$  is also symmetric w.r.t.  $\theta = \theta^*$ . As a result, a simple bisection search algorithm can efficiently find the global optimal solution  $\theta^*$  to (12). An example algorithm is given in Algorithm 1.

Note that condition (13) can be satisfied by a variety of energy fields. For example, for Laplacian field  $u(x) = \sqrt{\gamma}e^{-\gamma|x|}$ , we have  $\tau(t) = (1 + \gamma t)e^{-\gamma t}$ , and  $\tau'(t) = -\gamma^2 t e^{-\gamma t}$ ; for Gaussian field  $u(x) = (\frac{2\gamma}{\pi})^{\frac{1}{4}}e^{-\gamma x^2}$ , we have  $\tau(t) = e^{-\gamma t^2/2}$ ,

---

### Algorithm 1 Search for the optimal rotation angle

---

- 1) Let  $\theta_L = 0$  and  $\theta_R = \frac{\pi}{2}$ . Choose an integer  $T \geq 1$  for smoothing (for sampling noise tolerance).
  - 2) Let  $\theta_c = \frac{1}{2}(\theta_L + \theta_R)$ . Take uniformly  $T$  points in  $[\theta_L, \theta_c]$ , i.e.,  $\theta_i = \theta_c - \frac{i}{T}(\theta_c - \theta_L)$ , and compute  $\bar{\rho}_L(\theta_L, \theta_R) = \frac{1}{T} \sum_{i=1}^T \rho(\theta_i)$  using (11) and (12). Compute  $\bar{\rho}_R(\theta_L, \theta_R)$  in the similar way.
  - 3) If  $\bar{\rho}_L > \bar{\rho}_R$ , then  $\theta_R = \theta_c$ ; otherwise,  $\theta_L = \theta_c$ .
  - 4) Repeat from Step 2) until  $\theta_R - \theta_L$  small enough. Then  $\theta^* = \theta_c$  is found.
- 

and  $\tau'(t) = -\gamma t e^{-\gamma t^2/2}$ . In both cases, condition (13) is satisfied.

### C. Source Detection

In the coordinate system  $\mathcal{C}_\theta$  under optimal rotation  $\theta = \theta^*$  (assuming alignment on the  $x$ -axis), the left and right singular vectors of  $\hat{\mathbf{H}}_\theta$  can be modeled as  $\hat{\mathbf{u}}_1 = \frac{1}{2}(\mathbf{u}_1(\theta^*) + \mathbf{u}_2(\theta^*)) + \mathbf{e}_u$  and  $\hat{\mathbf{v}}_1 = \mathbf{v}_1(\theta^*) + \mathbf{e}_v$ , respectively. Correspondingly, the  $y$ -coordinates of the sources can be found using estimator (7) based on reflected correlation

$$\hat{y}_1^S(\hat{\mathbf{v}}_1; \theta^*) = \hat{y}_2^S(\hat{\mathbf{v}}_1; \theta^*) = \frac{1}{2} \underset{t \in \mathbb{R}}{\text{argmax}} \hat{R}(t; \hat{\mathbf{v}}_1). \quad (14)$$

To find the  $x$ -coordinates, note that the function  $u_1(x) = \frac{1}{2}(u(x - x_1^S) + u(x - x_2^S))$  is symmetric at  $x = \frac{1}{2}(x_1^S + x_2^S)$ . Therefore, the center of the two sources can be found by

$$\hat{c} = \frac{1}{2} \underset{t \in \mathbb{R}}{\text{argmax}} \hat{R}(t; \hat{\mathbf{u}}_1). \quad (15)$$

In addition, after estimating  $\hat{y}_1^S$ , the marginal power density function  $u(x)$  can be obtained as  $\hat{u}(y) = \hat{v}_1(y - \hat{y}_1^S)$ , where  $\hat{v}_1(y)$  is a regression function from  $\hat{\mathbf{v}}_1$  (for example, by linear interpolation among  $y_1, y_2, \dots, y_{n_c}$ ). As a result, the  $x$ -coordinates of the two sources can be found using similar techniques as spread spectrum early gate synchronization [14], and obtained as  $\hat{x}_1^S(\theta^*) = \hat{c} - \hat{d}$  and  $\hat{x}_2^S(\theta^*) = \hat{c} + \hat{d}$ , where

$$\hat{d} = \underset{d \geq 0}{\text{argmax}} Q(d; \hat{\mathbf{u}}_1, \hat{\mathbf{v}}_1) \quad (16)$$

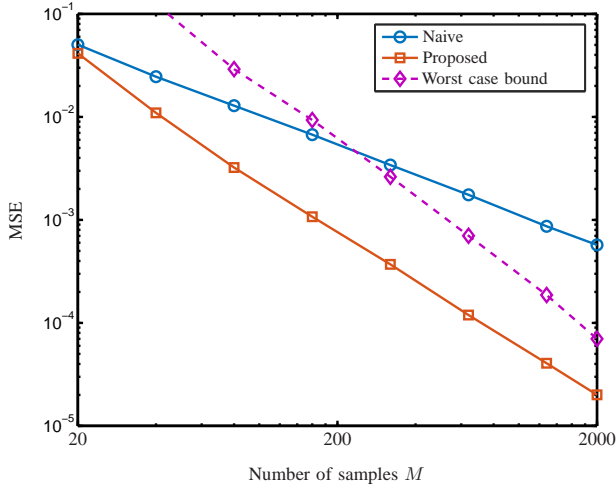
and

$$Q(d; \hat{\mathbf{u}}_1, \hat{\mathbf{v}}_1) \triangleq \frac{1}{2} \int_{-\infty}^{\infty} \hat{u}_1(x) (\hat{u}(x - \hat{c} - d) + \hat{u}(x - \hat{c} + d)) dx.$$

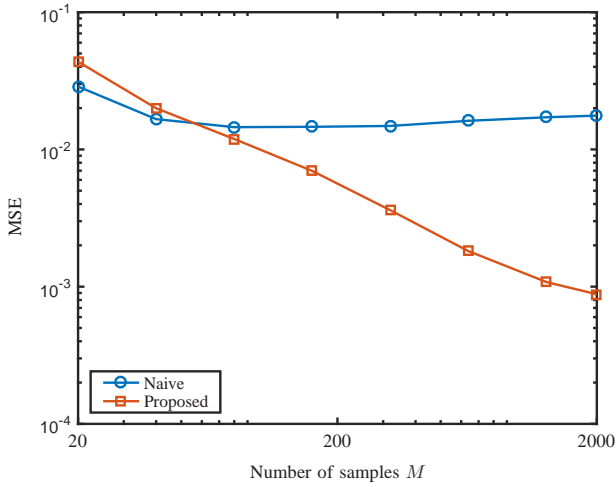
It is straight-forward to show that  $Q(d; \hat{\mathbf{u}}_1, \hat{\mathbf{v}}_1)$  is maximized at  $d^* = \frac{1}{2}|x_1^S - x_2^S|$ .

As a benchmark, consider a naive scheme that estimates  $x_1^S$  and  $x_2^S$  by analyzing the peaks of  $\hat{\mathbf{u}}_1$ . However, such naive strategy cannot work for small source separation, because if  $d = \frac{1}{2}|x_1^S - x_2^S|$  is too small, the aggregate power density function  $\tilde{u}_1(x) = u(x - x_1^S) + u(x - x_1^S - d)$  would be unimodal and there is only one peak in  $\hat{\mathbf{u}}_1$ . As a comparison, the proposed procedure estimator from procedure (14)–(16) does not such a limitation.





(a) One source



(b) Two sources

Figure 2. MSE of the source location versus the number of samples  $M$ .

## V. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed location estimator in both single source and double source cases. Two sources are placed in the area  $[-0.5, 0.5] \times [-0.5, 0.5]$  uniformly and independently at random, with the restriction that the distance between the two sources is no more than 0.5.<sup>3</sup> The power field generated by each source in an underwater environment is modeled as  $h_k(x, y) = e^{-20(x-x_k^s)^2 - 20(y-y_k^s)^2}$ ,  $k = 1, 2$ . There are  $M$  power measurements taken in the area  $\mathcal{A} = [-1, 1] \times [-1, 1]$  uniformly at random. The parameter  $n_c$  of the proposed observation matrix  $\hat{\mathbf{H}} \in \mathbb{R}^{n_c \times n_c}$  is chosen as the largest integer satisfying  $n_c(\log n_c)^2 \leq M/C$ , for  $C = 1$ .

As a benchmark, the proposed location estimation is compared with the naive scheme, which determines the source location directly from the position of the measurement sample

<sup>3</sup>When the two sources are far apart, the problem degenerates to two single-source-localization problems.

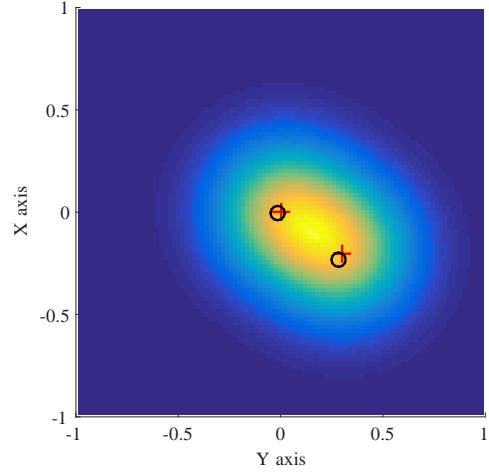


Figure 3. Localizing two sources using  $M = 200$  samples, where red crosses denote the true source locations, and black circles denote the estimates. The color map represents the aggregate power field generated by the two sources.

that observes the highest power. In the two source case, the naive algorithm aims at detecting either one of the sources, and the corresponding localization error is computed as  $\mathcal{E}_{\text{naive}}^2 = \min\{\|\hat{\mathbf{s}}_{\text{naive}} - \mathbf{s}_1\|^2, \|\hat{\mathbf{s}}_{\text{naive}} - \mathbf{s}_2\|^2\}$ . As a comparison, the localization error of the proposed algorithm is computed as  $\mathcal{E}^2 = \frac{1}{2}(\|\hat{\mathbf{s}}_1 - \mathbf{s}_1\|^2 + \|\hat{\mathbf{s}}_2 - \mathbf{s}_2\|^2)$ .

Fig. 2 depicts the MSE of the source location versus the number of samples  $M$ . In the single source case, the coefficient of the worst case upper bound (10) is chosen as  $C_\mu = 1$  to demonstrate the asymptotic decay rate of the worst case squared error bound. The decay rate of the analytic worst case error bound is roughly the same as the MSE obtained from the numerical experiment. It is expected that as  $M$  increases, the two curves merge in an asymptotic way. As a benchmark, the proposed scheme requires less than half of the samples to achieve similar performance to that of the naive baseline even for small  $M$  (around 50). More importantly, it demonstrates a higher MSE decay rate, where for medium  $M$  (around 200), the proposed scheme reduces the number of samples to 1/10. In the double source case, there is an error floor for the naive scheme, because the location that observes the highest power may not be either one of the source locations. As a comparison, there is no error floor in for proposed scheme as  $M$  increases.

Fig. 3 shows an example on simultaneously localizing two sources (red crosses). Although the aggregate power field has only one peak, the algorithm (black circles) is able to separate the two sources.

## VI. CONCLUSIONS

This paper developed source localization algorithms from a few power measurement samples, while no specific energy-decay model is assumed. Instead, the proposed method only exploited the structural property of the power field generated by the sources. Analytical results were developed to demonstrate that the proposed algorithm decreases the localization error at a higher rate than the baseline algorithm when the

number of samples increases. In addition, a rotated eigenstructure analysis technique was derived for simultaneously localizing two sources. Numerical results demonstrate the performance advantage in localizing single or double sources.

#### ACKNOWLEDGMENTS

This research was supported, in part, by National Science Foundation under Grant NSF CNS-1213128, CCF-1410009, CPS-1446901, Grant ONR N00014-15-1-2550, and Grant AFOSR FA9550-12-1-0215.

#### APPENDIX

##### A. Proof of Lemma 1

$$\begin{aligned}\tau'(t) &= \frac{d}{dt} \int_{-\infty}^{\infty} u(x)u(x-t)dx \\ &= \int_{-\infty}^{\infty} -u(x)u'(x-t)dx \\ &= -\int_{-\infty}^0 u(z+t)u'(z)dz - \int_0^{\infty} u(z+t)u'(z)dz \\ &= -\int_{-\infty}^0 u(z+t)u'(z)dz + \int_0^{\infty} u(z+t)u'(-z)dz \\ &= -\int_{-\infty}^0 u(z+t)u'(z)dz + \int_{-\infty}^0 u(-w+t)u'(w)dw \\ &= -\int_{-\infty}^0 [u(z+t) - u(-z+t)]u'(z)dz \\ &= -\int_{-\infty}^0 [u(z+t) - u(z-t)]u'(z)dz \\ &< 0\end{aligned}\tag{17}$$

(18)

$$\begin{aligned}&= -\int_{-\infty}^0 [u(z+t) - u(-z+t)]u'(z)dz \\ &= -\int_{-\infty}^0 [u(z+t) - u(z-t)]u'(z)dz \\ &< 0\end{aligned}\tag{19}$$

where (17) is due to the change of variable  $z = x - t$  and  $u'(z) = -u'(-z)$ , (18) is to change the variable  $z = -w$ , (19) exploits the fact that  $u(x) = u(-x)$ , and the last inequality is due to  $u(z+t) - u(z-t) > 0$  and  $u'(z) > 0$  for all  $z < 0$ .

##### B. Proof of Proposition 1 (Sketch)

Let  $e(x) \triangleq \hat{u}(x) - u(x - x_1^S)$ . Then  $\mathbb{E}\{e(x)\} = 0$  for all  $x$ , and  $\mathbb{E}\{\int_{-\infty}^{\infty} e(x)e(-x+t)dx\} = 0$  for  $t$  due to the zero mean and independent assumption on  $\mathbf{e}_u$ . Similarly,  $\mathbb{E}\{\int_{-\infty}^{\infty} u(x - x_1^S)e(-x+t)dx\} = 0$  for all  $t$ . As a result, we have

$$\begin{aligned}\mathbb{E}\{\hat{R}(t; \hat{\mathbf{u}}_1)\} &= \mathbb{E}\left\{\int_{-\infty}^{\infty} \left(u(x - x_1^S) + e(x)\right) \right. \\ &\quad \times \left(u(-x - x_1^S + t) + e(-x + t)\right) dx \Big\} \\ &= R(t; x_1^S)\end{aligned}$$

which is maximized at  $t = 2x_1^S$ .

On the other hand, it is easy to verify that  $\hat{R}(t + 2x_1^S; \hat{\mathbf{u}}_1)$  and  $\hat{R}(-t + 2x_1^S; \hat{\mathbf{u}}_1)$  has the same distribution, since the elements of  $\mathbf{e}_u$  are i.i.d.. Therefore,  $\mathbb{E}\{\arg \max_t \hat{R}(t; \hat{\mathbf{u}}_1)\} = \arg \max_t R(t; x_1^S) = x_1^S$ , which confirms that the estimator  $\hat{x}_1^S$  is unbiased.

##### C. Proof of Theorem 1

To simplify the algebra, we only focus on the dominant terms w.r.t.  $n$  as  $n$  goes large.

1) *Upper Bound of the Sampling Error:* For notational convenience, define  $u_1(x) = u(x - x_1^S)$  and  $v_1(y) = u(x - y_1^S)$ . Consider the sampling position  $(x, y) \in \mathcal{G}_{ij}$ . Using a Taylor expansion, we have

$$\begin{aligned}|h_1(x, y) - h_1(x_1, y_1)| &= \alpha |u_1(x)v_1(y) - u_1(x_1)v_1(y_1)| \\ &= \alpha |(u_1(x_1) + u_1'(x_1)(x - x_1)) \\ &\quad \times (v_1(y_1) + v_1'(y_1)(y - y_1)) - u_1(x_1)v_1(y_1) \\ &\quad + o(x - x_1) + o(y - y_1)| \\ &= \alpha |u_1(x_1)v_1'(y_1)(y - y_1) + v_1(y_1)u_1'(x_1)(x - x_1) \\ &\quad + o(x - x_1) + o(y - y_1)| \\ &\leq \alpha u(0)K_u \frac{L}{n_c} + o\left(\frac{L}{n_c}\right)\end{aligned}$$

from the property  $u(x) \leq u(0)$  and  $|u'(x)| \leq K_u$ .

From (2), we have

$$\begin{aligned}|\hat{H}_{ij} - H_{ij}| &= \left(\frac{L}{n_c}\right)^2 |h_1(x, y) - h_1(x_1, y_1)| \\ &\leq \alpha u(0)K_u \frac{L^3}{n_c^3} + o\left(\frac{L^3}{n_c^3}\right).\end{aligned}\tag{20}$$

As a result,

$$\begin{aligned}\|\mathcal{P}_{\Omega}(\hat{\mathbf{H}}_c - \mathbf{H})\|_F^2 &= \sum_{(i,j) \in \Omega} |\hat{H}_{ij} - H_{ij}|^2 \\ &\leq M (\alpha u(0)K_u L^3 / n_c^3)^2 \triangleq \epsilon^2.\end{aligned}$$

2) *Matrix Completion with Noise and Singular Vector Perturbation:* When there is sampling noise, the performance of matrix completion can be evaluated by the following result.

**Lemma 2** (Matrix completion with noise [13]). *Consider that  $\epsilon$  in (5) is chosen such that  $\|\mathcal{P}_{\Omega}(\hat{\mathbf{H}} - \mathbf{H})\|_F \leq \epsilon$ . Then, with high probability,*

$$\delta \triangleq \|\hat{\mathbf{H}}_c - \mathbf{H}\|_F \leq 4\sqrt{\frac{(2+p)n_c}{p}}\epsilon + 2\epsilon\tag{21}$$

where  $p = M/n_c^2$ .

As we focus on not too small  $n_c$ , which is chosen to be such that  $M \approx Cn_c(\log n_c)^2$ , the bound (21) can be simplified as

$$\begin{aligned}\delta &\leq 4\sqrt{\frac{(2 + Cn_c(\log n_c)^2/n_c^2)n_c}{Cn_c(\log n_c)^2/n_c^2}}\epsilon + 2\epsilon \\ &\approx \sqrt{\frac{32}{C}} \frac{n_c}{\log n_c} \epsilon.\end{aligned}$$

Let  $\mathbf{u}_1$  and  $\hat{\mathbf{u}}_1 = \mathbf{u}_1 + \mathbf{e}_u$  be the dominant left singular vectors of  $\mathbf{H}$  and  $\hat{\mathbf{H}}_c$ , respectively. We exploit the following classical result from singular vector perturbation analysis.

**Lemma 3** (Singular vector perturbation [15]). *Let  $\sigma_1$  and  $\sigma_2$  be the first and second dominant singular values of  $\mathbf{H}$ . Then,*

$$\sin \angle(\mathbf{u}_1, \hat{\mathbf{u}}_1) \leq \frac{2\|\hat{\mathbf{H}}_c - \mathbf{H}\|_F}{\sigma_1 - \sigma_2}.$$

By exploiting Lemma 3 for our case, we have

$$\begin{aligned} \sin \angle(\mathbf{u}_1, \hat{\mathbf{u}}_1) &= \sqrt{1 - |\mathbf{u}_1^T(\mathbf{u}_1 + \mathbf{e}_u)|^2} \\ &= \sqrt{-2\mathbf{u}_1^T\mathbf{e}_u + |\mathbf{u}_1^T\mathbf{e}_u|^2} \\ &\approx \sqrt{2|\mathbf{u}_1^T\mathbf{e}_u|} \end{aligned}$$

where  $|\cdot|$  denotes the absolute value operator, and we drop the second order term  $|\mathbf{u}_1^T\mathbf{e}_u|^2$ , since  $|\mathbf{u}_1^T\mathbf{e}_u|$  is small as we focus on large  $n_c(M)$ . We also note that  $\mathbf{u}_1^T\mathbf{e}_u \leq 0$ .

Consider that we have chosen  $M \approx Cn_c(\log n_c)^2$ , and moreover,  $\mathbf{H}$  is a rank-1 matrix with singular value  $\sigma_1 = \alpha$ . As a result,

$$\begin{aligned} 2|\mathbf{u}_1^T\mathbf{e}_1| &\approx \sin^2 \angle(\mathbf{u}_1, \hat{\mathbf{u}}_1) \leq \left(\frac{2\delta}{\alpha}\right)^2 \leq \frac{4}{\alpha^2} \frac{32}{C} \left(\frac{n_c}{\log n_c}\right)^2 \epsilon^2 \\ &= \mu_u L^6 n_c^{-3} \end{aligned}$$

where  $\mu_u = 128u(0)^2 K_u^2$ .

3) *Estimator based on Reflected Correlation:* Let  $e(x) = \hat{u}(x) - u_1(x)$ . Note that by construction,  $e(x)$  is an linear interpolation of the error vector  $\mathbf{e}_u$  at  $x = x_1, x_2, \dots, x_{n_c}$ . Let  $\mathbf{e}_{-t}$  be an  $n$ -dimensional vector that takes value  $\mathbf{e}_{-t}(i) = e(-x_i + t)$ . From the i.i.d. assumption of  $\mathbf{e}_u$ , the two vectors are identical in distribution, i.e.,  $\mathbf{e}_u \stackrel{d}{=} \mathbf{e}_{-t}$ . Therefore, based on assumptions A1 and A2, we can make the following approximation

$$\int_{-\infty}^{\infty} u_1(x)e(-x+t)dx \approx \mathbf{u}_1^T\mathbf{e}_{-t} \stackrel{d}{=} \mathbf{u}_1^T\mathbf{e}_u.$$

Moreover, we have

$$\int_{-\infty}^{\infty} e(x)e(-x+t)dx \approx \mathbf{e}_u^T\mathbf{e}_{-t} \approx 0$$

since we focus on not too small  $n_c(M)$ , and the elements of  $\mathbf{e}_u$  (and  $\mathbf{e}_{-t}$ ) are zero mean and independent.

Define a reflected correlation function as

$$R(t; x_1^S) = \int_{-\infty}^{\infty} u(x - x_1^S)u(-x - x_1^S + t)dx.$$

Then, it follows that  $R(t; x_1^S) = \tau(2x_1^S - t)$ . As a result, we have

$$\begin{aligned} \hat{R}(t; \hat{\mathbf{u}}_1) &= \int_{-\infty}^{\infty} (u_1(x) + e(x)) (u_1(-x+t) + e(-x+t)) dx \\ &= \int_{-\infty}^{\infty} u_1(x)u_1(-x+t)dx + \int_{-\infty}^{\infty} u_1(x)e(-x+t)dx \\ &\quad + \int_{-\infty}^{\infty} e(x)u_1(-x+t)dx + \int_{-\infty}^{\infty} e(x)e(-x+t)dx \\ &\approx R(t; x_1^S) + \mathbf{u}_1^T\mathbf{e}_{-t} + \mathbf{e}_u^T\mathbf{u}_{-t} \end{aligned}$$

and hence  $R(t; x_1^S) - \hat{R}(t; \hat{\mathbf{u}}_1) \approx -\mathbf{u}_1^T\mathbf{e}_{-t} - \mathbf{e}_u^T\mathbf{u}_{-t}$ .

Recall that  $\hat{t} = 2\hat{x}_1^S$  maximizes  $\hat{R}(\hat{t}; \hat{\mathbf{u}}_1)$  and  $t^* = 2x_1^S$  maximizes  $R(t^*; x_1^S) = \tau(2x_1^S - t^*)$ . We have

$$\begin{aligned} R(t^*; x_1^S) - \hat{R}(\hat{t}; \hat{\mathbf{u}}_1) &= \tau(0) - \tau(2|\hat{x}_1^S - x_1^S|) \\ &\leq 2|\mathbf{u}_1^T\mathbf{e}_1| \\ &\leq \mu_u L^6 n_c^{-3} + o(n_c^{-3}) \end{aligned}$$

where  $o(n_c^{-3})$  is due to the fact that we keep omitting the higher order terms. Finally, we obtain

$$\tau(2|\hat{x}_1^S - x_1^S|) = 1 - \mu_u L^6 n_c^{-3} + o(n_c^{-3})$$

and hence,

$$|\hat{x}_1^S - x_1^S| \leq \frac{1}{2}\tau^{-1}(1 - \mu_u L^6 n_c^{-3} + o(n_c^{-3})).$$

#### D. Proof of Theorem 2

We first study the singular vectors in double source case.

**Lemma 4** (Singular vectors in two source case). *Let  $\mathbf{u}_k(\theta)$  and  $\mathbf{v}_k(\theta)$  be the vectors defined following (3) and (4) in the rotated coordinate system  $\mathcal{C}_\theta$ . The SVD of  $\mathbf{H}_\theta$  is given by*

$$\mathbf{H}_\theta = \alpha_1 \mathbf{p}_1 \mathbf{q}_1^T + \alpha_2 \mathbf{p}_2 \mathbf{q}_2^T \quad (22)$$

where  $\alpha_1 = \frac{\alpha}{2}\|\mathbf{u}_1 + \mathbf{u}_2\|\|\mathbf{v}_1 + \mathbf{v}_2\|$  and  $\alpha_2 = \frac{\alpha}{2}\|\mathbf{u}_1 - \mathbf{u}_2\|\|\mathbf{v}_1 - \mathbf{v}_2\|$  are the singular values, and

$$\begin{aligned} \mathbf{p}_1 &= \frac{\mathbf{u}_1 + \mathbf{u}_2}{\|\mathbf{u}_1 + \mathbf{u}_2\|}, & \mathbf{q}_1 &= \frac{\mathbf{v}_1 + \mathbf{v}_2}{\|\mathbf{v}_1 + \mathbf{v}_2\|} \\ \mathbf{p}_2 &= \frac{\mathbf{u}_1 - \mathbf{u}_2}{\|\mathbf{u}_1 - \mathbf{u}_2\|}, & \mathbf{q}_2 &= \frac{\mathbf{v}_1 - \mathbf{v}_2}{\|\mathbf{v}_1 - \mathbf{v}_2\|} \end{aligned}$$

are the corresponding singular vectors.

*Proof:* First,

$$\begin{aligned} \mathbf{H}_\theta &= \bar{\alpha}(\mathbf{u}_1 \mathbf{v}_1^T + \mathbf{u}_2 \mathbf{v}_2^T) \\ &= \frac{\bar{\alpha}}{2}[(\mathbf{u}_1 + \mathbf{u}_2)(\mathbf{v}_1 + \mathbf{v}_2)^T + (\mathbf{u}_1 - \mathbf{u}_2)(\mathbf{v}_1 - \mathbf{v}_2)^T] \\ &= \alpha_1 \mathbf{p}_1 \mathbf{q}_1^T + \alpha_2 \mathbf{p}_2 \mathbf{q}_2^T \end{aligned}$$

Hence, these four vectors form a decomposition of  $\mathbf{H}_\theta$ .

Second, we have

$$\begin{aligned} \mathbf{p}_1^T \mathbf{p}_2 &= c(\mathbf{u}_1 + \mathbf{u}_2)^T(\mathbf{u}_1 - \mathbf{u}_2) \\ &= c(\|\mathbf{u}_1\|^2 - \|\mathbf{u}_2\|^2) \\ &= 0 \end{aligned}$$

where  $c = 1/(\|\mathbf{u}_1 + \mathbf{u}_2\|\|\mathbf{u}_1 - \mathbf{u}_2\|)$ . Similarly,  $\mathbf{q}_1^T \mathbf{q}_2 = 0$ . In addition, all the four vectors have unit norm.

As a result, (22) is the SVD of  $\mathbf{H}_\theta$ .  $\blacksquare$

Consider an arbitrary coordinate system. W.l.o.g. (due to Assumption 1), assume that the first source is located at the origin,  $x_1^S = 0$  and  $y_1^S = 0$ , and the second source is away from the first source with distance  $D$  and angle  $\theta$  to the  $x$ -axis,  $x_2^S = D \cos \theta$  and  $y_2^S = D \sin \theta$ . In addition, defining

$$u_c(x, \theta) \triangleq u(x - D \cos \theta), \quad u_s(x, \theta) \triangleq u(x - D \sin \theta)$$

we have

$$\begin{aligned}\mathbf{u}_1 &= \sqrt{\delta} [u(x_1), u(x_2), \dots, u(x_N)]^T \\ \mathbf{v}_1 &= \sqrt{\delta} [u(y_1), u(y_2), \dots, u(y_M)]^T \\ \mathbf{u}_2 &= \sqrt{\delta} [u_c(x_1, \theta), u_c(x_2, \theta), \dots, u_c(x_N, \theta)]^T \\ \mathbf{v}_2 &= \sqrt{\delta} [u_s(y_1, \theta), u_s(y_2, \theta), \dots, u_s(y_M, \theta)]^T.\end{aligned}$$

Based on assumption A1 and A2, we have

$$\begin{aligned}\|\mathbf{u}_k\|^2 &= \left(\frac{L}{n}\right)^2 \sum_{i=1}^N u(x_i - x_k^S)^2 \approx \int_{x_1}^{x_{n-1}} u(x - x_k^S)^2 dx \\ &\approx \int_{-\infty}^{\infty} u(x - x_k^S)^2 dx = 1\end{aligned}\quad (23)$$

and similar integrals apply to  $\mathbf{v}_k$ .

As an equivalent statement to Theorem 2, we need to show that  $\rho(\theta)$  is a strictly increasing function in  $\theta \in (0, \frac{\pi}{4})$ . Equivalently, we should prove that the function

$$\begin{aligned}&\frac{\lambda_2(\mathbf{H}_\theta)^2}{\lambda_1(\mathbf{H}_\theta)^2} \\ &\approx \frac{\int_{-\infty}^{\infty} (u(x) - u_c(x, \theta))^2 dx \int_{-\infty}^{\infty} (u(x) - u_s(x, \theta))^2 dx}{\int_{-\infty}^{\infty} (u(x) + u_c(x, \theta))^2 dx \int_{-\infty}^{\infty} (u(x) + u_s(x, \theta))^2 dx} \\ &\triangleq \mu(\theta)\end{aligned}$$

is strictly increasing in  $\theta \in (0, \frac{\pi}{4})$ , where the approximated integrals are obtained from (23).

To simplify the notation, define the integration operator  $\langle \cdot \rangle$  as

$$\langle f \rangle \triangleq \int_{-\infty}^{\infty} f(x, \theta) dx$$

for a function  $f(x, \theta)$ . By definition, the integration operator is linear and satisfies the additive property, i.e.,  $\langle af \rangle = a \langle f \rangle$  and  $\langle f + g \rangle = \langle f \rangle + \langle g \rangle$ , for a constant  $a$  and a function  $g(x, \theta)$ . As a result,  $\langle (u - u_c)^2 \rangle = \langle u^2 \rangle + \langle u_c^2 \rangle - 2 \langle u \cdot u_c \rangle = 2(1 - \langle u \cdot u_c \rangle)$ , and the function  $\mu(\theta)$  can be written as

$$\mu(\theta) = \frac{(1 - \langle u \cdot u_c \rangle)(1 - \langle u \cdot u_s \rangle)}{(1 + \langle u \cdot u_c \rangle)(1 + \langle u \cdot u_s \rangle)}. \quad (24)$$

In addition, from the properties in calculus, if  $f(x, \theta)$  and  $\frac{\partial}{\partial \theta} f(x, \theta)$  are continuous in  $\theta$ , then

$$\begin{aligned}\frac{d}{d\theta} \langle f \rangle &= \frac{d}{d\theta} \int_{-\infty}^{\infty} f(x, \theta) dx \\ &= \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x, \theta) dx = \left\langle \frac{\partial}{\partial \theta} f \right\rangle.\end{aligned}$$

Therefore, defining

$$\begin{aligned}u'_c(x, \theta) &\triangleq \frac{d}{dx} u(x) \big|_{x=x-D \cos \theta} \\ u'_s(x, \theta) &\triangleq \frac{d}{dx} u(x) \big|_{x=x-D \sin \theta}\end{aligned}$$

we have

$$\begin{aligned}\frac{d}{d\theta} \langle u \cdot u_c \rangle &= \langle u \cdot \frac{\partial}{\partial \theta} u_c(x, \theta) \rangle = \langle u \cdot u'_c \rangle D \sin \theta \\ \frac{d}{d\theta} \langle u \cdot u_s \rangle &= \langle u \cdot \frac{\partial}{\partial \theta} u_s(x, \theta) \rangle = -\langle u \cdot u'_s \rangle D \cos \theta.\end{aligned}$$

With some algebra, the derivative of  $\mu(\theta)$  can be obtained as

$$\begin{aligned}\frac{d}{d\theta} \mu(\theta) &= \eta \left[ D \cos \theta \langle u \cdot u'_s \rangle (1 - \langle u \cdot u_c \rangle^2) \right. \\ &\quad \left. - D \sin \theta \langle u \cdot u'_c \rangle (1 - \langle u \cdot u_s \rangle^2) \right] \\ &= \eta \left[ -t \cdot \tau'(s) (1 - \tau(t)^2) + s \cdot \tau'(t) (1 - \tau(s)^2) \right]\end{aligned}$$

where  $\eta = 2(1 + \langle u \cdot u_c \rangle)^{-2} (1 + \langle u \cdot u_s \rangle)^{-2}$ ,  $t = D \cos \theta$ , and  $s = D \sin \theta$ .

Note that  $0 < s < t$  for  $0 < \theta < \frac{\pi}{4}$ . Applying condition (13), we have

$$\begin{aligned}\frac{d}{d\theta} \mu(\theta) &> \eta \cdot t \cdot \tau'(s) \left[ (1 - \tau(s)^2) - (1 - \tau(t)^2) \right] \\ &= \eta \cdot t \cdot \tau'(s) (\tau(t)^2 - \tau(s)^2) \\ &> 0\end{aligned}$$

since  $\tau'(s) < 0$  and  $\tau(t) < \tau(s)$  for  $0 < s < t$ .

This confirms that  $\mu(\theta)$  is a strictly increasing function, and hence  $\rho(\theta)$  is a strictly increasing function in  $\theta \in (0, \frac{\pi}{4})$ . The results in Theorem 2 is confirmed.

## REFERENCES

- [1] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1770–1778, 2008.
- [2] H.-D. Qi, N. Xiu, and X. Yuan, "A lagrangian dual approach to the single-source localization problem," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3815–3826, 2013.
- [3] X. Sheng and Y.-H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 44–53, 2005.
- [4] C. Meesookho, U. Mitra, and S. Narayanan, "On energy-based acoustic source localization for sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 365–377, 2008.
- [5] Y. Liu, Y. H. Hu, and Q. Pan, "Distributed, robust acoustic source localization in a wireless sensor network," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4350–4359, 2012.
- [6] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1553–1560, 1988.
- [7] R. Lefort, G. Real, and A. Drémeau, "Direct regressions for underwater acoustic source localization in fluctuating oceans," *Applied Acoustics*, vol. 116, pp. 303–310, 2017.
- [8] X. Nguyen, M. I. Jordan, and B. Sinopoli, "A kernel-based learning approach to ad hoc sensor network localization," *ACM Trans. on Sensor Networks*, vol. 1, no. 1, pp. 134–152, 2005.
- [9] Y. Jin, W.-S. Soh, and W.-C. Wong, "Indoor localization with channel impulse response based fingerprint and nonparametric regression," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1120–1127, 2010.
- [10] W. Kim, J. Park, J. Yoo, H. J. Kim, and C. G. Park, "Target localization using ensemble support vector regression in wireless sensor networks," *IEEE Trans. on Cybernetics*, vol. 43, no. 4, pp. 1189–1198, 2013.
- [11] S. Choudhary and U. Mitra, "Analysis of target detection via matrix completion," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2015, pp. 3771–3775.
- [12] S. Choudhary, N. Kumar, S. Narayanan, and U. Mitra, "Active target localization using low-rank matrix completion and unimodal regression," *arXiv preprint arXiv:1601.07254*, 2016.
- [13] E. J. Candes and Y. Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.
- [14] R. Peterson, R. Ziemer, and D. Borth, *Introduction to spread spectrum systems*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [15] V. Vu, "Singular vectors under random perturbation," *Random Structures & Algorithms*, vol. 39, no. 4, pp. 526–538, 2011.